

A Monte Carlo Time-Dependent Variational Principle

F. W. G. Transchel,^{*} A. Milsted,[†] and Tobias J. Osborne[‡]
Institut für Theoretische Physik, Appelstr. 2, Hannover, D-30167, Germany
 (Dated: November 21, 2014)

We generalize the Time-Dependent Variational Principle (TDVP) to dissipative systems using Monte Carlo methods, allowing the application of existing variational classes for pure states, such as Matrix Product States (MPS), to the simulation of Lindblad master equation dynamics. The key step is to use sampling to approximately solve the Fokker-Planck equation derived from the Lindblad generators. An important computational advantage of this method, compared to other variational approaches to mixed state dynamics, is that it is “embarrassingly parallel”.

Quantum many body systems are hard to solve. Even the largest supercomputers are stumped by the general case because the dimension of Hilbert space scales exponentially with the number of particles. However, most of this vast space corresponds to states that are highly *entangled*, a property which is *not* possessed by a great many physically relevant states [1–3]. Fortunately, this can be exploited by working with *variational classes* of states that parameterize this highly relevant corner of Hilbert space. In the case of one-dimensional systems, this is achieved by the hugely successful Density Matrix Renormalization Group (DMRG) technique [4–6], which can be understood [7] as a variational method based on Matrix Product States (MPS) — a class for which entanglement is upper-bounded by the dimension of the MPS parameter space [8]. Recently, the Time-Dependent Variational Principle (TDVP) has also been applied to MPS, providing a very promising framework for finding ground states, the simulation of dynamics, and for probing the excitation spectrum [9–11].

Although these advances have revolutionized the simulation of many body systems, most work so far has focused on unitary dynamics. Deviations from unitarity are generally from undesired couplings to the environment that destroy coherence and are, to that extent, merely an experimental nuisance. However, it has become increasingly clear that introducing dissipation in a controlled way by engineering the system-environment coupling can be instrumental in performing a number of very useful tasks in quantum information processing [12, 13]. From this perspective, being able to efficiently simulate dissipative dynamics is highly desirable.

Existing methods for the simulation of general dissipative dynamics are based either on Monte Carlo sampling or on variational classes of mixed states. On one hand, sampling wavefunctions [14, 15] is an easily parallelizable problem that is, however, in general limited to small systems due to the computational demands of evolving wavefunctions for each sample trajectory. On the other, general variational methods such as the TDVP for mixed states [16] suffer from the lack of a unique measure of information distance for comparing mixed states. More specific techniques, such as those using MPS on purified states with time-evolving block decimation [17, 18]

are somewhat less easy to parallelize. A natural way of combining these two approaches is to choose a variational class of pure states to represent the wavefunction components of a density matrix, using Monte Carlo sampling to simulate dissipation. This was suggested in [17] and applied with a mean field wavefunction ansatz in [19], yet the use of more complicated variational classes in this way has so far focused on the specific case of generating approximations of thermal states [20, 21].

In this Letter, we introduce a general variational approach to dissipative dynamics that makes use of an arbitrary pure state variational class, thus avoiding the need to explicitly fix a measure of information distance for mixed states. We do this by applying Monte Carlo sampling to the TDVP flow equations derived from the Lindblad master equation describing the system. The use of sampling creates an effective variational class of mixed states from the pure state ansatz and has the added benefit of making the resulting algorithm “embarrassingly parallel” [22]. We implement the method for MPS and trial it on a simple spin-chain system with nearest-neighbor interactions and spin-flip dissipation to check convergence of the sampling. We then test it on a larger XXZ Heisenberg chain driven at the edges, a system with known analytic solutions.

We begin in a general dissipative setting where the dynamics of a state ρ , belonging to a Hilbert space \mathcal{H} , is given by a Lindblad master equation of the form

$$\partial_t \rho = -i[K, \rho] - \frac{1}{2} \sum_{\alpha} (L_{\alpha}^{\dagger} L_{\alpha} \rho + \rho L_{\alpha}^{\dagger} L_{\alpha} - 2L_{\alpha} \rho L_{\alpha}^{\dagger}), \quad (1)$$

where K is Hermitian, L_{α} are bounded operators and we can write the RHS as $Q\rho + \rho Q^{\dagger} + \sum_{\alpha} L_{\alpha} \rho L_{\alpha}^{\dagger}$ using $Q = -iK - \frac{1}{2} \sum_{\alpha} L_{\alpha}^{\dagger} L_{\alpha}$. We approximate a general mixed state at a time t as

$$\rho_t = \int_{\mathcal{M}} p_t(\bar{a}, a) |\Psi(a)\rangle \langle \Psi(a)| d\bar{a} da, \quad (2)$$

where $|\Psi(a)\rangle$ belongs to a variational class of wavefunctions with parameters a , such as MPS for one-dimensional lattice systems, and $p_t(\bar{a}, a)$ is a probability distribution over those parameters.

If we come back to the integral from before, note that it is over the submanifold of Hilbert space $\mathcal{M} \in \mathcal{H}$ formed

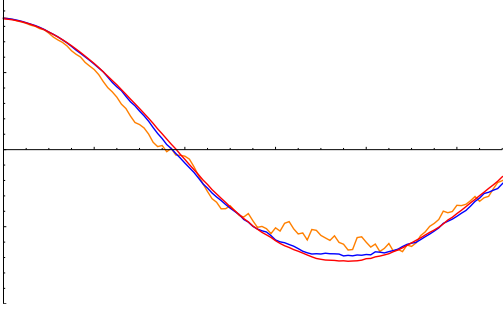


Fig. 1. Comparison of the accuracy of an ensemble of MPS representations at 480 (yellow), 4800 (blue), 48000 (red) samples. The exact solution is not shown because the difference from the red curve would not be visible on this plot. We plot from $t = 80$ because the difference between the curves is too small to be seen for $t < 80$.

by $|\Psi(a)\rangle \in \mathcal{M}$ and cannot generally be evaluated efficiently since the dimension of \mathcal{M} is, despite being much smaller than $\dim(\mathcal{H})$, still large. However, as we show, this ansatz lends itself naturally to Monte Carlo sampling. In the following, we assume for reasons of simplicity and without loss of generality, that $|\Psi(a)\rangle$ depends holomorphically on a .

Applying the master equation (1) to the ansatz (2) gives us an infinitesimal time step $\partial_t \rho_t$ in terms of the evolution $\partial_t p_t$ of the probability distribution together with the action of the operators Q and L_α on the pure state $|\Psi(a)\rangle$. If we wish to maintain the form (2) whilst evolving the state, the latter must be approximated by a vector $|\Phi(b)\rangle \equiv b^j \partial_{a^j} |\Psi(a)\rangle$ in the tangent space \mathbb{T}_a to the variational manifold \mathcal{M} at point a (note that repeated indices are summed over unless otherwise stated). This kind of approximation forms the basis of the time-dependent variational principle (TDVP), as explained further in the supplementary material. The tangent vector parameters b_Q optimally approximating $Q|\Psi(a)\rangle$ can be found via $b_Q \approx \arg \min_{b'} |Q|\Psi(a)\rangle - |\Phi(b')\rangle|^2$, with b_α for each L_α following in identical fashion. After partial integration (discarding surface terms), we obtain an effective master equation

$$\partial_t \rho_t = \int_{\mathcal{M}} \left[-\partial_{a^j} (p_t b_Q^j) - \text{c.c.} + \partial_{a^k} \partial_{\bar{a}^l} (p_t b_\alpha^k \bar{b}_\alpha^l) \right] |\Psi(a)\rangle \langle \Psi(a)| da d\bar{a}, \quad (3)$$

which evolves states only within the mixed state ansatz (2), approximating the exact Lindblad dynamics in a locally optimal (in time) way. We now use sampling to evaluate the integral by taking the variational parameters a to be random variables, which may lead to a solution (see [23]) in form of a stochastic differential equation

(SDE)

$$da^j(t) = \left(b_Q^j + \langle \bar{L}_\alpha \rangle b_\alpha^j \right) dt + b_\alpha^j dw_\alpha(t). \quad (4)$$

These are identical to the TDVP flow equations [9] with additional dissipative noise captured by the b_α and can be sampled from via numerical integration beginning from some starting parameters a_0 . There exist other ways to unravel a master equation, including quantum jumps (see [24], [25]). However the Quantum State Diffusion method [15] is best suited to the TDVP. If the variational manifold \mathcal{M} captures the full Hilbert space, the presented method reduces to the exact QSD. Expectation values for ρ_t can be computed approximately with the standard convergence of $1/\sqrt{N}$, where N is the number of samples.

Numerical integration can be performed for each sample using, for example, the following algorithm implementing the Euler method:

1. Generate a starting state for the sample by initializing the variational parameters a^j with suitable values.
2. To evolve from time t to $t+dt$, evaluate (4) and set $a^j(t+dt) = a^j(t) + da^j(t)$.
3. Normalize $|\Psi(a)\rangle$ if necessary and restore a canonical form for a^j as needed.
4. Calculate expectation values of interest, i.e. energy or magnetization of the lattice or its elements and go to step 2. How one would calculate such ensemble expectations is also explained in the supplementary material.

The need for normalization and the restoration of a canonical form varies depending on the chosen pure state variational class. For example, in the case of MPS the first is needed and, due to redundancy in the choice of parameters, the second is recommended for numerical conditioning.

Since samples are completely independent of one another they can be evaluated simultaneously using any sufficiently capable computer processors, making this method “embarrassingly parallel” with regard to scaling in N . This is fortunate, since the variance of approximate expectation values scales with $\frac{1}{\sqrt{N}}$, such that we would need to square the number of sample runs to double the accuracy. For our tests, we thought about sending phishing mails to fellow physicists to take over their personal computers as computational zombie nodes, but were ultimately content with the existing computation servers available for use at our institute.

We implemented the method using matrix product states (MPS) as the pure state variational ansatz, which is a class well-suited to approximating states of many-body systems in one dimension, where the Hilbert space

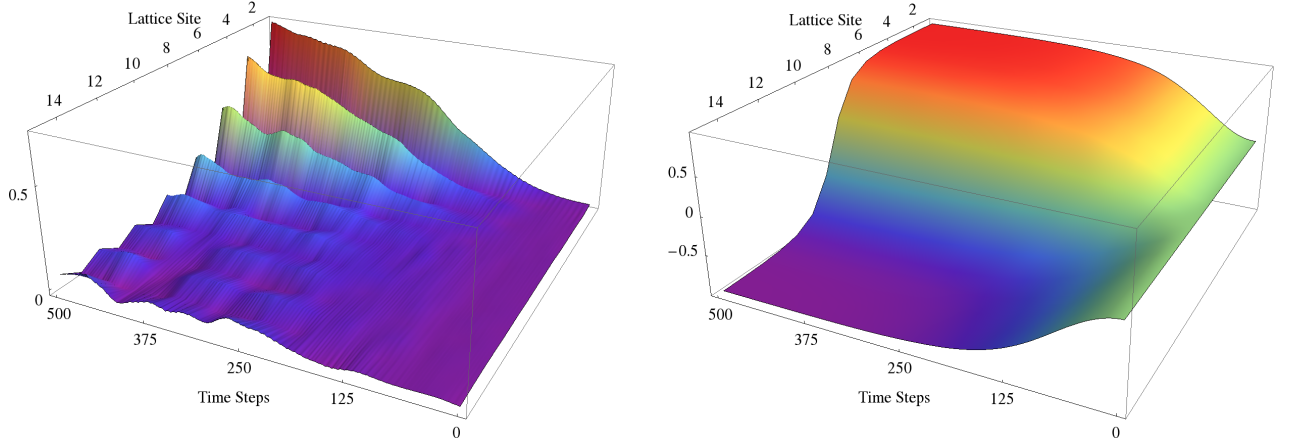


Fig. 2. **Left** Time-resolved two-point $\langle \sigma^x \sigma^x \rangle$ -correlation function (vertical axis) of site 8 with all other sites for the Heisenberg-type model K_{XZ} (see text) under spin-flip dissipation $L = \sum_n \sigma_n^+$, simulated using $n = 16$ sites, $N = 1000$ samples and bond dimensions $D = 32$. We clearly observe that the spin correlations comply with the expected antiferromagnetic domain behaviour of the K_{XZ} model, while the absolute values of the spin expectations are skewed due to the inherent symmetry breaking induced by the finite length of the lattice. **Right** Same plot for bihomogeneous dissipation of the form $L_{\alpha=1 \dots \frac{n}{2}} = \sigma_{\alpha}^+$, $L_{\alpha=(\frac{n}{2}+1) \dots n} = \sigma_{\alpha}^-$ on a Heisenberg XZ lattice with $n = 16$, $N = 1000$ and $D = 64$. Correlations are smoothed throughout the lattice due to the fact that the spin alignment is mediated between the two antipodal domains. Note that this was not at all clear at the domain flip region in the center of the lattice.

is taken to be isomorphic to $\mathcal{H} = (\mathbb{C}^d)^{\otimes N_{\text{sites}}}$, with a local dimension of d . MPS for finite N_{sites} with open boundary conditions have the form $|\Psi[A(t)]\rangle = \sum_{s=1}^d v_L^\dagger A_1^{s_1} \dots A_N^{s_N} v_r |s\rangle$, where $|s\rangle = |s_1 \dots s_N\rangle$, $A_n^s \in M_D(\mathbb{C})$ and v_L^\dagger, v_r are boundary vectors that may be absorbed into A_1 and A_N respectively. The corresponding variational manifold $\mathcal{M}_{\text{MPS}(D)}$ constitutes a submanifold of \mathcal{H} , where D is the so-called bond-dimension — the maximum Schmidt rank of the Schmidt-decompositions made by cutting between any two neighboring sites. We implement the method by extending evoMPS [26], an open-source implementation of the TDVP for MPS, calculating the MPS tangent vectors approximating local Q (consisting of, say, nearest-neighbor terms) in the same way as for a local Hamiltonian [9]. It is sufficient for the following examples to restrict L_α to on-site operators. Since applying an on-site operator to an MPS results in a tangent vector $(L_\alpha |\Psi[A]\rangle \in \mathbb{T}_{[A]})$, no approximation is needed in calculating the b_α .

To test the method we compare it to an exact calculation of dissipative dynamics on a two-qubit ($N_{\text{sites}} = 2$) lattice $\mathcal{H} = \mathbb{C}^2 \otimes \mathbb{C}^2$. We define single site Lindblad operators $L_\alpha = \sigma_\alpha^+$ for $\alpha = 1 \dots N_{\text{sites}}$ and a Hamiltonian part $K_{XZ} = \sum_{n=1}^{N_{\text{sites}}-1} \sigma_n^x \sigma_{n+1}^x + \lambda \sigma_n^z \sigma_{n+1}^z$ and simulate the dissipative dynamics, beginning with a highly entangled MPS (within the constraints of the chosen bond dimension). We can readily check that a full state tomography of the two-site density matrix differs from the exact analytic solution only by the expected variance of $\frac{1}{\sqrt{N}}$, where N is the number of sample paths calculated. Figure 1 shows convergence with increasing N . We are also able to replicate Rabi oscillations between the spin com-

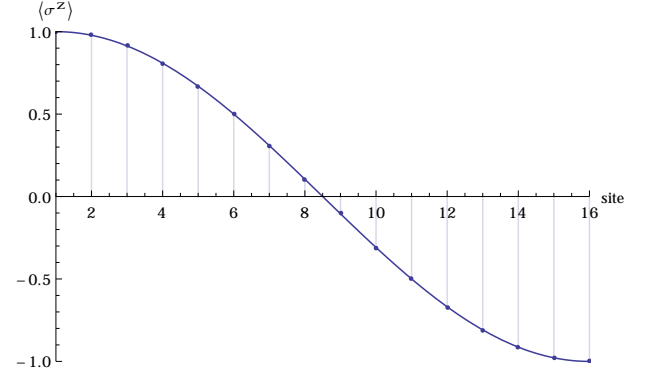


Fig. 3. Numerical results for the XXZ chain with edge driving ($n_{\text{sites}} = 16, \epsilon = 1, \lambda = 1, D = 24$), showing good agreement with the analytical solution in [27] even for a very low number of samples (300).

ponents of the two sites in the x direction. We conclude that our method and implementation are technically and numerically robust enough to be trialed on larger systems, beyond the reach of exact numerics.

We simulated the same dynamics as above on a longer spin chain with $N_{\text{sites}} = 16$, finding that the clustering of correlations grows as we increase the bond dimension D to match the number of degrees of freedom the system needs to to approximate the dynamics to good accuracy. Figure 2 shows the time evolution of the σ^x two-point correlation functions (using the 8th site as the reference) for bond dimensions $D = 2$ and $D = 32$. Correlation flow is somewhat comparable to Rabi oscillations in a 2-qubit system.

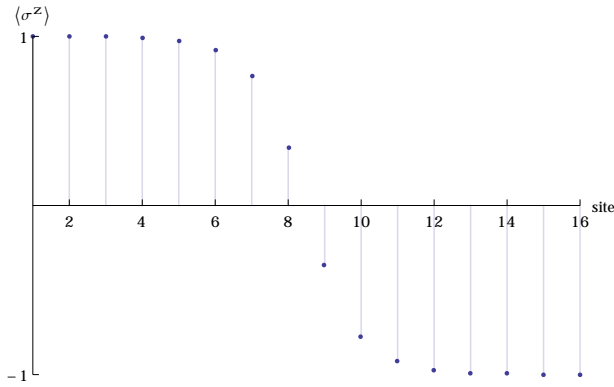


Fig. 4. Bihomogenous dissipation for the XXZ chain at low interaction strength (or strong driving, $\epsilon = 10^{-3}$, $\lambda = 1$, $n = 16$, $N = 2500$, $D = 64$)

An XXZ Heisenberg model of the form $K = \sum_{n=1}^{N_{\text{sites}}-1} \epsilon [2\sigma_n^+ \sigma_{n+1}^- + 2\sigma_n^- \sigma_{n+1}^+ + \lambda \sigma_n^z \sigma_{n+1}^z]$ (with ϵ controlling the relative strength of the site coupling with respect to the dissipation) has been analytically solved for strong driving in [27]. It is thus suitable for demonstrating the method on larger system sizes where the use of a pure state variational ansatz like MPS leads to large advantages over other Monte Carlo methods. When dissipation is introduced by Lindblad operators $L_1 = \sigma_1^+$, $L_2 = \sigma_N^-$ acting on the ends of the lattice, we find (see figure 3) that our method can replicate analytic results at small lattice sizes at the expense of high sample numbers. For larger N however, we clearly observe errors made due to the chosen bond dimension becoming a limiting factor. In these cases the information about dissipation at the far ends of the system does not permeate to the center, where we observe the largest deviation from analytic results — the center spins don't appear to be coupled to the environment at all.

It should be noted that convergence is expected to be slow for this model because the dynamics are critical [28]. The information about the center of the chain propagates only slowly to the reservoir at the edges and thus renders the XXZ chain with edge pumping in some sense the worst kind of system one could imagine for this method. With that in mind it seems remarkable that the results are at least qualitatively comparable to the analytics.

We then explored non-integrable systems like an XXZ chain with bihomogenous dissipation consisting of a Lindblad operator for each site $L_n = \sigma_n^+$ for $n \leq N/2$ and $L_m = \sigma_m^-$ for $m > N/2$. For large interaction strength, the results are intuitive as there are two clearly separate domains of magnetization in the system, while for weak interactions (see figure 6), especially near the center, we can see interference between the up and down pumping of magnetization. Furthermore we conclude that the bond dimension of the MPS need not approach the $2^{N/2}$ needed to represent an arbitrary state exactly in order to capture

interference effects and maintain sufficient amounts of entanglement throughout the lattice. This system behaves like a one-dimensional bar magnet in the sense that we are able to tune the system parameters ϵ and λ in such a way that we can explore the behavior at the center, i.e. the zone of spin domain change. For large magnetic interaction and weak dissipative coupling we find that the method is able to highlight interference patterns near the center.

In this paper we presented a Monte Carlo extension to the TDVP, expanding the scope of its applicability. In the future it seems natural to apply the method to larger dissipative systems. That includes, but is not limited to, larger lattice sizes as well as larger internal Hilbert space sizes. Since the method still is, in essence, a variational approach, it can outperform exact diagonalization while errors stay limited as long as we can afford to produce enough samples. With the inherent statistical nature of the method, it should be possible to approximate solutions to great accuracy where analytic solutions are highly non-tractable, as long as parallel computing resources are available. It is also worth noting that the inherent scalability of the method would allow to efficiently use it on even larger grid scales such as supercomputers to gain insights into previously unreachable domains. The software developed for this work is based on the evoMPS project [26] and is freely available under an open source license.

This work was supported by the ERC grants QFTCMPS, and SIQS by the cluster of excellence EXC 201 Quantum Engineering and Space-Time Research. F. W. G. T. wishes to acknowledge useful discussions with R. F. Werner, K. Abdelkhalek and B. Neukirchen.

* fabian.transchel@itp.uni-hannover.de

† ashley.milsted@itp.uni-hannover.de

‡ tobias.osborne@itp.uni-hannover.de

- [1] J. Eisert, M. Cramer, and M. B. Plenio, *Rev. Mod. Phys.* **82**, 277 (2010).
- [2] M. B. Hastings, *J. Stat. Mech. Theor. Exp.* **2007**, P08024 (2007).
- [3] T. J. Osborne, *Phys. Rev. Lett.* **97**, 157202 (2006).
- [4] S. R. White, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [5] S. R. White, *Phys. Rev. B* **48**, 10345 (1993).
- [6] I. Peschel, X. Want, M. Kaulke, and K. Hallberg, eds., *Density-Matrix Renormalization, a New Numerical Method in Physics*, Lecture Notes in Physics, Berlin Springer Verlag, Vol. 528 (1999).
- [7] U. Schollwöck, *Ann. Phys.* **326**, 96 (2011).
- [8] M. Fannes, B. Nachtergaele, and R. Werner, *Commun Math Phys* **144**, 443 (1992); J. I. Cirac and F. Verstraete, *J. Phys. A - Math. Theor.* **42**, 504004 (2009).
- [9] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pizorn, H. Verschelde, and F. Verstraete, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [10] J. Haegeman, B. Pirvu, D. J. Weir, J. I. Cirac, T. J.

- Osborne, H. Verschelde, and F. Verstraete, *Phys. Rev. B* **85**, 100408 (2012).
- [11] A. Milsted, J. Haegeman, T. J. Osborne, and F. Verstraete, arXiv:1207.0691 (2012).
- [12] J. T. Barreiro, M. Mueller, P. Schindler, D. Nigg, T. Monz, M. Chwalla, M. Hennrich, C. F. Roos, P. Zoller, and R. Blatt, *Nature* **470**, 486 (2011); H. Krauter, C. A. Muschik, K. Jensen, W. Wasilewski, J. M. Petersen, J. I. Cirac, and E. S. Polzik, *Phys. Rev. Lett.* **107**, 80503 (2011).
- [13] S. Diehl, A. Micheli, A. Kantian, B. Kraus, H. P. Büchler, and P. Zoller, *Nat. Phys.* **4**, 878 (2008); B. Kraus, H. P. Büchler, S. Diehl, A. Kantian, A. Micheli, and P. Zoller, *Phys. Rev. A* **78**, 42307 (2008); F. Verstraete, M. M. Wolf, and J. Ignacio Cirac, *Nat. Phys.* **5**, 633 (2009); J. Eisert and T. Prosen, arXiv:1012.5013 (2010); M. J. Kastoryano, F. Reiter, and A. S. Sørensen, *Phys. Rev. Lett.* **106**, 90502 (2011); C. A. Muschik, E. S. Polzik, and J. I. Cirac, *Phys. Rev. A* **83**, 52312 (2011); F. Pastawski, L. Clemente, and J. I. Cirac, *Phys. Rev. A* **83**, 12304 (2011); K. G. H. Vollbrecht, C. A. Muschik, and J. I. Cirac, *Phys. Rev. Lett.* **107**, 120502 (2011); A. Mari and J. Eisert, *Phys. Rev. Lett.* **108**, 120602 (2012).
- [14] C. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics* (Springer, 2004).
- [15] N. Gisin and I. C. Percival, *J. Phys. A: Math. Gen.* **25**, 5677 (1992).
- [16] C. V. Kraus and T. J. Osborne, *Phys. Rev. A* **86**, 5 (2012).
- [17] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [18] M. Zwolak and G. Vidal, *Phys. Rev. Lett.* **93**, 207205 (2004).
- [19] H. Pichler, J. Schachenmayer, A. J. Daley, and P. Zoller, *Phys. Rev. A* **87**, 033606 (2013).
- [20] S. R. White, *Phys. Rev. Lett.* **102**, 190601 (2009); E. M. Stoudenmire and S. R. White, *New J. Phys.* **12**, 055026 (2010).
- [21] S. Garnerone, arXiv:1309.0851 (2013).
- [22] B. J. Breen, C. E. Weidert, J. F. Lindner, L. M. Walker, K. Kelly, and E. Heidtmann, *Phys. Rev. Lett.* **76**, 347 (2008).
- [23] C. W. Gardiner, *Handbook of stochastic methods : for physics, chemistry and the natural sciences*, Springer series in synergetics, 13 (Springer, 2002).
- [24] A. J. Daley, (2014), arXiv:1405.6694 [quant-ph].
- [25] L. Bonnes, D. Charrier, and A. M. Läuchli, *Phys. Rev. A* **90**, 033612 (2014).
- [26] A. Milsted, M. Lewerenz, and F. W. G. Transchel, <https://github.com/amilsted/evoMPS> (2012).
- [27] T. Prosen, *Phys. Rev. Lett.* **107**, 137201 (2011).
- [28] T. Hikihara and A. Furusaki, *Phys. Rev. B* **69**, 064427 (2004).
- [29] P. E. Kloeden, E. Platen, and H. Schurz, *Numerical solution of SDE through computer experiments*, corr. 3. print. ed., Universitext (Springer, Berlin, 2003) pp. XIV, 292 p.

Supplementary material

Time Dependent Variational Principle

Given a starting state $|\Psi(a)\rangle$ belonging to a pure state variational class with parameters a and a differential equation describing some dynamics, we wish to compute the approximate time evolution of the state with the restriction that it must remain in the variational class.

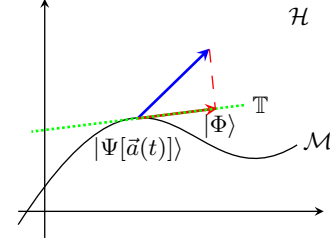


Fig. 5. TDVP schematic showing the projection of a tangent vector (blue arrow) in the full Hilbert space \mathcal{H} onto the tangent space \mathbb{T} of a variational sub-manifold \mathcal{M} , resulting in a tangent vector $|\Phi\rangle$ (red arrow).

We assume that an exact infinitesimal time step has the form $|\Psi(t + dt)\rangle = |\Psi(a(t))\rangle + dt O|\Psi(a(t))\rangle$, with some operator $O \in \mathcal{B}(\mathcal{H})$, which would be the Hamiltonian (times i) of the system in the case of unitary dynamics. As illustrated in Figure 5, the locally optimal approximation to the dynamics is then given by the vector $|\Phi(b)\rangle \equiv b^j \partial_{a^j} |\Psi(a)\rangle$, belonging to the tangent plane \mathbb{T} to the variational manifold \mathcal{M} at the point a , which best approximates the vector $O|\Psi(a(t))\rangle$ (blue arrow):

$$b = \arg \min_{b'} ||\Phi(b')\rangle - O|\Psi(a(t))\rangle|^2.$$

Solving this optimization problem is equivalent to numerical integration of a set of flow equations for a

$$\partial_t a^j = -g^{jk} \partial_{a^k} \langle \Psi(a) | O | \Psi(a) \rangle,$$

where g^{jk} is the inverse of the “metric” $g_{jk}(\bar{a}, a) \equiv \partial_{\bar{a}^j} \partial_{a^k} \langle \Psi(a) | \Psi(a) \rangle$. In the case where g_{jk} is not invertible, a pseudo-inverse may be used instead. The TDVP flow equations can also be derived via the principle of least action. For more details, see for example [9] (supplementary material).

Derivation of the Fokker-Planck equation for sampling from the Lindblad master equation

As summarized in the main text, we want to capture the behavior of systems of the form

$$\begin{aligned} \partial_t \rho &= -i[K, \rho] - \frac{1}{2} \sum_{\alpha} (L_{\alpha}^{\dagger} L_{\alpha} \rho + \rho L_{\alpha}^{\dagger} L_{\alpha} - 2L_{\alpha} \rho L_{\alpha}^{\dagger}) \\ &= Q\rho + \rho Q^{\dagger} + \sum_{\alpha} L_{\alpha} \rho L_{\alpha}^{\dagger}, \end{aligned} \quad (5)$$

with $Q = -iK - \frac{1}{2} \sum_{\alpha} L_{\alpha}^{\dagger} L_{\alpha}$, where K is Hermitian and ρ the state that is to be evolved. L_{α} is the set of arbitrary Lindblad operators that model dissipation.

$$\partial_t \rho_t = \int_{\mathcal{M}} \partial_t p_t(\bar{a}, a) |\Psi(a)\rangle \langle \Psi(a)| da d\bar{a}, \quad (6)$$

where $p_t(\bar{a}, a)$ is a time-dependent probability distribution over the pure state variational parameters a and \mathcal{M} is the sub manifold of Hilbert space formed by the states in the variational class. Since this integral cannot be (efficiently) performed for a general class of pure states, we exploit stochastic calculus to sample from it.

First, we review some details of stochastic differential equations (SDE). Following Gardiner [23], the expectation value of a function of a random variable described by an Ito SDE is

$$\begin{aligned} \left\langle \frac{df[x(t)]}{dt} \right\rangle &= \frac{d}{dt} \langle f[x(t)] \rangle \\ &= \left\langle r[x(t), t] \partial_x f + \frac{1}{2} u[x(t), t]^2 \partial_x^2 f \right\rangle, \end{aligned} \quad (7)$$

where $x(t)$ is a random variable, $r(x, t)$ is the drift coefficient, and $u(x, t)$ is the diffusion coefficient. If x takes values according to the conditional probability density $p(x, t|x_0, t_0)$, for initial conditions x_0 at t_0 , then we can write

$$\begin{aligned} \frac{d}{dt} \langle f[x(t)] \rangle &= \int dx f(x) \partial_t p(x, t|x_0, t_0) \\ &= \int dx \left[r(x, t) \partial_x f + \frac{1}{2} u(x, t)^2 \partial_x^2 f \right] p(x, t|x_0, t_0). \end{aligned} \quad (8)$$

If we now integrate by parts and discard surface terms, we get

$$\begin{aligned} \int dx f(x) \partial_t p &= \int dx f(x) \left\{ -\partial_x [r(x, t)p] \right. \\ &\quad \left. + \frac{1}{2} \partial_x^2 [u(x, t)^2 p] \right\}. \end{aligned} \quad (9)$$

Choosing $f(x) = 1$, we learn that

$$\partial_t p = -\partial_x [r(x, t)p] + \frac{1}{2} \partial_x^2 [u(x, t)^2 p]. \quad (10)$$

It should thus be clear that the evolution of p is governed by the drift and diffusion coefficients r and u .

In the case of many complex variables $x^j \in \mathbb{C}$, an Ito SDE may take the form

$$dx^j = r^j(\bar{x}, x, t) dt + U_{\alpha}^j(\bar{x}, x, t) dw_{\alpha}, \quad (11)$$

where $dw_{\alpha} = \frac{1}{\sqrt{2}}(du_{\alpha} + i dv_{\alpha})$ are complex Wiener processes constructed from real Wiener processes du_{α}, dv_{α}

such that $\langle dw_{\alpha} d\bar{w}_{\beta} \rangle = \delta_{\alpha\beta} dt$ and $\langle dw_{\alpha} dw_{\beta} \rangle = 0$. One can then derive

$$\begin{aligned} \int d\bar{x} dx f(\bar{x}, x) \partial_t p &= \\ \int d\bar{x} dx f(\bar{x}, x) \left\{ -\partial_{x^j} [r^j(\bar{x}, x, t)p] - \text{c.c.} \right. \\ &\quad \left. + \partial_{x^k} \partial_{\bar{x}^l} [U_{\alpha}^k(\bar{x}, x, t) \overline{U_{\alpha}^l(\bar{x}, x, t)} p] \right\}, \end{aligned} \quad (12)$$

where $r^j, U_{\alpha}^j \in \mathbb{C}$ and U is now called the diffusion matrix. Again, we can easily obtain the evolution of $p(\bar{x}, x, t)$ by setting $f(\bar{x}, x) = 1$.

We may attempt to find such an equation for p_t in (6), viewing the entries of ρ as functions of complex random variables a with expectation values calculated by integrating over \mathcal{M} . In fact, the TDVP delivers exactly the drift vector and the diffusion matrix needed. Inserting (6) into (5) and using the TDVP to approximate the vectors $Q|\Psi(a)\rangle$ and $L_{\alpha}|\Psi(a)\rangle$ as $b_Q^j \partial_{a^j} |\Psi(a)\rangle$ and $b_{\alpha}^j \partial_{a^j} |\Psi(a)\rangle$ respectively, we arrive (after partial integration, discarding surface terms) at

$$\begin{aligned} \partial_t \rho_t &= \int_{\mathcal{M}} \left[-\partial_{a^j} (p_t b_Q^j) - \text{c.c.} + \partial_{a^k} \partial_{\bar{a}^l} (b_{\alpha}^k \overline{b_{\alpha}^l} p_t) \right] \\ &\quad \times |\Psi\rangle \langle \Psi| da d\bar{a}, \end{aligned}$$

the RHS of which has the same form as (12). We can thus read off a Fokker-Plank equation for p_t

$$\partial_t p_t = -\partial_{a^j} (p_t b_Q^j) - \text{c.c.} + \partial_{a^k} \partial_{\bar{a}^l} (b_{\alpha}^k \overline{b_{\alpha}^l} p_t). \quad (13)$$

and obtain an Ito SDE for the variational parameters a

$$da^j(t) = \left(b_Q^j + \langle \bar{L}_{\alpha} \rangle b_{\alpha}^j \right) dt + b_{\alpha}^j dw_{\alpha}(t). \quad (14)$$

Interpreting a as a random, stochastic variable, one could have naïvely chosen an ansatz of the form $a = b_Q dt + b_{\alpha} dw_{\alpha}$, where $dw_{\alpha}(t)$ is a complex Wiener process (white noise), corresponding to the linear expression $d(|\psi\rangle) = Q|\psi\rangle dt + L_{\alpha}|\psi\rangle dw_{\alpha}$. This however is unsuccessful, since the differential Ito calculus for $d(|\psi\rangle \langle \psi|)$ resolves to

$$d(|\psi\rangle \langle \psi|) = |d\psi\rangle \langle \psi| + |\psi\rangle \langle d\psi| + L_{\alpha}|\psi\rangle \langle \psi| \bar{L}_{\alpha} dt \quad (15)$$

and we actually want to evolve the system as a probability density functional instead of a linear Hilbert space vector. By integrating eq. (14) we can evolve a pure state component $|\Psi(a)\rangle$ of some initial ρ such that it samples the evolution of the full mixed state. Using N such samples, properties of ρ_t can be approximated with an error (variance) that scales as $1/\sqrt{N}$.

How to sample from a Wiener process

It is vital for understanding the presented Monte Carlo scheme that we do not attempt to correctly approximate

the actual mixed state ρ at some time t , but rather the expectation values of observables of interest. For example, the observable σ^z transforms under the time evolution $\rho(t) = e^{t\mathcal{L}}(\rho_0)$, where \mathcal{L} is the completely-positive trace-preserving map corresponding to the RHS of (5), as

$$\langle \sigma^z \rangle(t) = \text{tr}(\sigma^z e^{t\mathcal{L}}(\rho_0)) \quad (16)$$

and the stochastic expectation value for N samples is

$$\langle \langle \sigma^z \rangle \rangle(t) = \frac{1}{N} \sum_{l=1}^N \langle \sigma^z \rangle_l(t), \quad (17)$$

where $\langle \sigma_j^z \rangle_l$ is the expectation value of σ_j^z for the l th pure state sample.

Wiener processes were introduced as a tool to analyze and explain Brownian motion statistically. For example Kloeden [29] defines a Wiener process $W = W(t | t > 0)$ to be a continuous Gaussian process with independent increments such that

$$W(0) = 0 \text{ a.s.}, \quad \langle W(t) \rangle = 0, \\ \text{and } \text{Var}(W(t) - W(s)) = t - s.$$

It can be discretized into n steps with length Δt as $W_{n\Delta t} = \sqrt{\Delta t} \sum_{i=1}^n X_i$, where X_i are Gaussian random variables. In this form numerical treatment is easy as

long as good Gaussian pseudo-random numbers are available. In our case we use the NumPy Python framework incorporating the Mersenne Twister MT 19937 algorithm, which is widely used for Monte Carlo calculations.

Supplementary plots

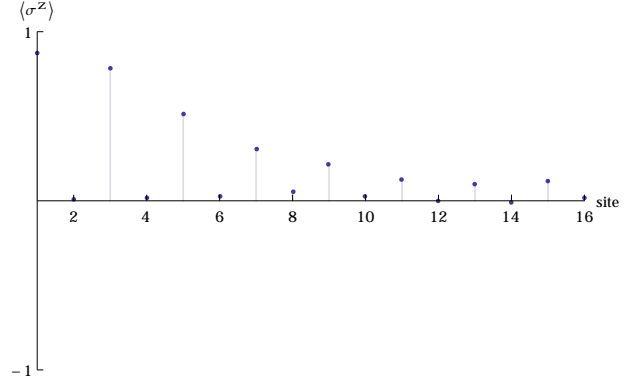


Fig. 6. Homogenous dissipation of the form $L_\alpha = \sigma_\alpha^+$ on the Heisenberg KX chain with $(n = 16, N = 1000, D = 32)$. The expectation values are skewed because the first spin points up due to initial driving and the last down due to anti-ferromagnetic constraints. Note how the absolute values do not become negative because the dissipation is strictly positive.